THE POTENTIAL DISTRIBUTION OF *ZOROTYPUS HUBBARDI* CAUDELL (ZORAPTERA: ZOROTYPIDAE) IN NORTH AMERICA, AS PREDICTED BY ECOLOGICAL NICHE MODELING

Ismael A. Hinojosa-Díaz, Elisa Bonaccorso, and Michael S. Engel

(IAH, MSE) Division of Entomology, Natural History Museum, 1460 Jayhawk Boulevard, Snow Hall, and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045-7523, U.S.A. (e-mail: hinojosa@ku.edu); (EB) Division of Ornithology, Natural History Museum, 1345 Jayhawk Boulevard, Dyche Hall, and Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045-7163, U.S.A.

Abstract.—Owing to their minute sizes, frail wings, and low abundances in entomological collections, zorapterans have been assumed to be highly endemic and poor dispersers. Wide distributions of some species are thought to be induced by human activity. Herein, we use the Genetic Algorithm for Rule-set Prediction (GARP) to model the potential distribution of *Zorotypus hubbardi* Caudell, one of 32 living species of the order Zoraptera. According to the model, the potential distribution of *Z. hubbardi* covers the southeastern U.S., with western limits in eastern Texas, eastern Oklahoma, and southeastern Kansas, and northern limits in southeastern Iowa in the West and southernmost New York in the East; the model's predictions exclude *Z. hubbardi* from the Appalachian Mountains. The northern and western boundaries of the potential distribution show that the ecological conditions for the natural occurrence of the species are present, and that the presence of colonies would not be solely dependent on coexistence with humans. New feral records from the areas shown in the potential distribution would establish the certainty of the model.

Key Words: ecological niche, *Zorotypus*, Polyneoptera, predictive modeling, zorapteran

Zorotypus hubbardi Caudell is one of two species of Zoraptera occurring in the United States. This order of minute (2–3.5 mm in total length), soft-bodied insects, contains a single family with 32 living species restricted principally to tropical latitudes; only four species occur north of the Tropic of Cancer (two in the United States, and two in Tibet). The low diversity of zorapteran species contrasts with the overwhelming diversity of insects as a whole; however, there are phylogenetic and biological aspects that make this group of insects worthy of study. The monophyly of Zoraptera is supported by several characters, among them a unique pattern of wing venation, dimerous tarsi, and mating via "mating hook" (Engel 2003a, b, 2004). However, the phylogenetic placement of the order among hemimetabolous insects is controversial. In the best-supported hypoth-

esis, Zoraptera is a sister group to Embiodea (a.k.a., webspinners) (Engel and Grimaldi 2000, Grimaldi and Engel 2005).

These enigmatic insects live gregariously under the bark of decaying logs; they feed principally on fungal hyphae and spores, but also prey on nematodes or minute arthropods. Two distinct morphs (sometimes referred to as "castes," which should not be confused with the true castes of social insects) are present within each species and in the same colony. Blind, wingless individuals, which are more common during the life of the colony, and winged-eyed, dispersive individuals, which are generally rare in the colony; the latter individuals are produced when resources become depleted and the colony becomes crowded. Dispersive individuals shed their wings after arriving at a new log and found new colonies. The northernmost limit of the distribution of *Z. hubbardi* (in south-central and southeastern U.S.), is likely the result of human activities, as most colonies are found in sawdust piles (rather than natural logs), relying on the warmth of the decaying material that is available for just a few seasons. As the sawdust pile becomes unsuitable for the colony, dispersal is required; as a result, the northernmost colonies become extinct and subsequent reintroductions are required to maintain the presence of the species (Engel 2003a). Formerly, zorapterans were thought to have restricted dispersal capabilities making them highly endemic, but wide distributions (as for *Z. hubbardi*) and occurrence on oceanic islands, suggest a different scenario. Herein, we asses the potential distribution of *Z. hubbardi*, which in turn will serve to formulate hypotheses about the natural or introduced condition of the northernmost colonies, as well as ecological factors necessary for the persistence of the species in specific areas.

METHODS

Occurrence points and environmental data.—We based our ecological niche models on 79 occurrence points of *Z. hubbardi* in its native range. Occurrence points were obtained from literature records, and georeferenced to the nearest minute using GEOLocate (Rios and Bart 2005). Environmental coverages used in the process included 24 electronic maps summarizing aspects of topography (U.S. Geological Survey's Hydro-1K data set http://edcdaac.usgs.gov/gtopo30/hydro/) and climate (http://biogeo.berkeley.edu/worldclim/down.htm). Because most occurrence points dated from the 1950's, we used a dataset of climatic coverages for the time period 1930–1960. All coverages were resampled at $0.1°$ pixel resolution and clipped to eastern North America for model development. Compilation of environmental datasets and visualization of the models was achieved using Arc View (ESRI 2002).

Ecological niche modeling.—To develop ecological niche models, we used the Genetic Algorithm Rule-set Prediction (GARP) (Stockwell and Noble 1992, Stockwell 1999, Stockwell and Peters 1999). This approach focuses on modeling ecological niches; the conjunction of ecological conditions within a species is able to maintain populations without immigration (Grinnell 1917, MacArthur 1972). GARP relates ecological characteristics of known occurrence points to those of points sampled from the rest of the study region, developing a series of decision rules that best summarize those factors associated with the species presence (Feria and Peterson 2002).

Occurrence points are first divided into training and test data sets (Stockwell and Peters 1999). GARP, then applies an iterative process of rule selection, evaluation, testing, and incorporation or rejection (Stockwell and

Peters 1999). It chooses a method from a set of possibilities (e.g., logistic regression, bioclimatic rules) applied to the training data, and a rule is developed or evolved; rules may evolve by a number of means (e.g., truncation, point changes, crossing-over) to maximize predictivity. Predictive accuracy (for intrinsic use in the model refinement) is then evaluated based on 1250 points sampled randomly from the study region as a whole (Stockwell and Peters 1999). The change in predictive accuracy from one iteration to the next is used to evaluate whether a particular rule should be incorporated into the model, and the algorithm runs either 1000 iterations or until convergence occurs (Stockwell and Peters 1999).

Extensive testing has demonstrated GARP's ability to predict ecological and geographic distributions of species in diverse ecological, geographic, and taxonomic contexts (Anderson et al. 2003; Peterson and Cohoon 1999; Peterson 2001; Stockwell and Peterson 2002, 2003). Herein, we used a desktop version of GARP (http://www.lifemapper.org/desktopgarp).

To reduce the environmental data set to coverages providing the highest predictivity accuracy, we performed a variety of jackknife manipulations (Peterson et al. 2003). We ran multiple iterations of models omitting coverages, or sets of coverages, systematically. Then, we examined the correlation between the inclusion and exclusion of each coverage and omission error, and removed those coverages that had a strong detrimental contribution to model quality (when $r > 0.05$).

Ecological niche models developed with GARP can be projected onto landscapes via spatial queries to detect conditions fitting those modeled as the species' niche. Projection of models onto the landscape from which input occurrence points were taken, indicates geographic distribution of suitable conditions (i.e., a potential geographic distribution for the species). Those models can also be projected onto alternate landscapes in space (Peterson and Vieglais 2001) and time (Rice et al. 2003). Herein, we projected the model for the native distribution of the species back onto eastern North America to estimate the potential distribution of the species in the whole region.

Since GARP uses a "random-walk" approach and predictions vary among runs, we generated 100 models for the species, and selected the 10 "best" models (those falling in an optimal combination of error measures; Anderson et al. 2003). These 10 models were summed to obtain maps for the potential distribution of Z. hubbardi across eastern North America.

Testing ecological niche models.—The ecological niche models generated were tested as follows. Available distributional points were divided onto either side of the median latitude and median longitude. Two of the quadrants (northwest and southeast) were used to develop models in order to predict the distributions of points in the other two quadrants (southwest and northeast) and *vice versa*. Then, the proportional area predicted present × number of extrinsic data points was used as a random expectation of successful prediction of points if no non-random association existed between prediction and test points; a chi-square approach (1 df) was used to test the significance of the departure from random expectations (Peterson and Shaw 2003). This test represents a challenge for model predictivity because it asseses the ability of models to predict into broad, unsampled regions.

RESULTS

According to the jackknife analyses, all topographic layers were highly corre-
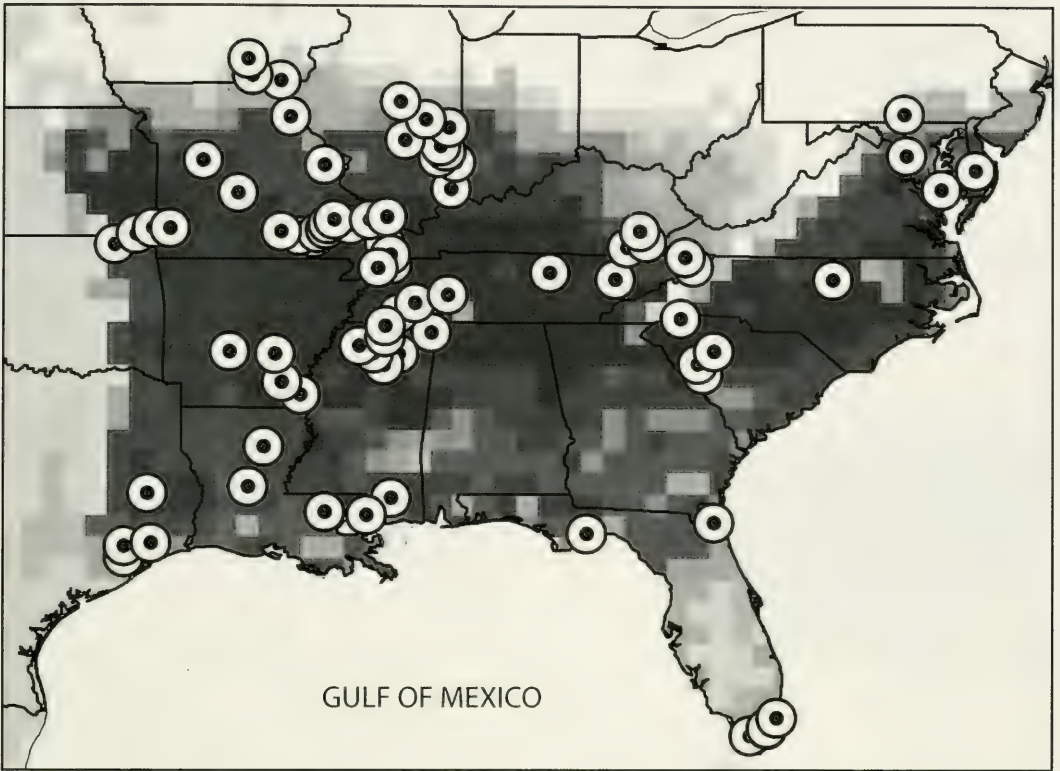
Fig. 1.   Prediction of the potential distribution of *Zorotypus hubbardi* in eastern North America. Dots show occurrence data on which ecological niche models were based. Darker shading indicates greater model agreement in predicting potential presence.

lated with high omission error (all correlations $r > 0.05$); therefore, they were excluded from further manipulations. After all analyses, seven environmental layers (frequency of ground frost, annual precipitation, frequency of wet days, and annual minimum, mean, and maximum temperature) were chosen to develop the final ecological niche models.

The models developed to predict the potential distribution of *Z. hubbardi* identified broad areas throughout the southeastern United States, extending centrally to the eastern extreme of Texas (particularly bordering the Gulf Coast), eastern half of Oklahoma, and southeastern Kansas. The northern limits reach southeastern Iowa in the West

Table 1.   Summary of statistical tests of model accuracy in the quadrant test.

Test[1]	Proportional Area Predicted Present[2]	N	Expected Correct	Expected Incorrect	Observed Correct	Observed Incorrect	X^2	P
NS predicts SN	30%	42	13	29	19	23	4.13	0.04
SN predicts NS	17%	37	6	31	11	26	4.48	0.03

[1] Two of the quadrants (northwest and southeast, NS) predict the distributions of points in the other two quadrants (southwest and northeast, SN) and vice versa.

[2] Area predicted present by all 10 models generated/total area of analysis.
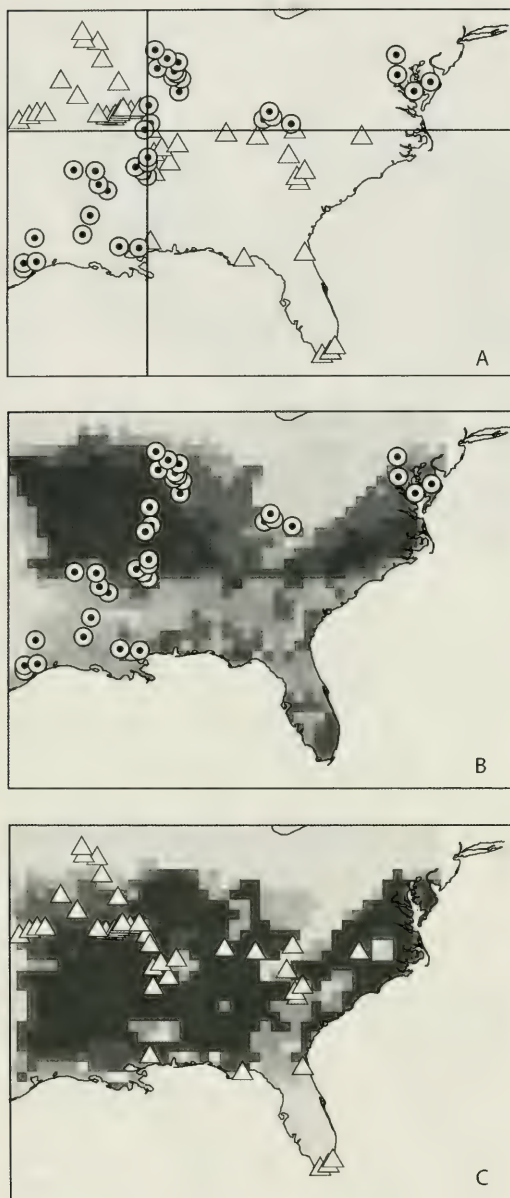
Fig. 2. Illustration of the quadrant test used to assess model predictivity for *Zorotypus hubbardi*. In panel A, we show the available distributional points for the species, divided in two groups (northwest and southeast, southwest and northeast). In panel B, points in quadrants northwest and southeast predict the distribution of points in quadrants southwest and northeast. Panel C shows the reciprocal test.

and the southernmost areas of New York in the East. The model excludes the Appalachian Mountains (Fig. 1).

The ecological niche models were highly predictive of the potential distribution of *Z. hubbardi*, based on the quadrant test of model predictivity (Table 1). In Fig. 2, the quadrant test is illustrated: 42 and 37 points were available in the two pairs of quadrants. In the first test, based on area predicted present, 12 of 42 points were expected to be predicted present correctly; however, 20 were predicted correctly (chi-square test, $P < 0.05$). The reciprocal test had an expected correct prediction of 6 of 37 points, but 11 of 37 points were correctly predicted (chi-square test, $P < 0.05$).

Visualization of patterns of distribution of species in ecological space provides a fertile ground for thinking about why species are distributed in some areas and not in others (Peterson and Shaw 2003). As an example, Fig. 3 shows the ecological space available across the species range in terms of Mean Annual Temperature and Mean Annual Precipitation, and Mean Annual Temperature and Frequency of Wet Days (light gray diamonds), as well as the ecological space occupied by the species according to its potential distribution (dark gray squares). In both cases, *Z. hubbardi* is present in zones with intermediate values for these three variables. It would be interesting to compare this pattern with other *Zorotypus* species to study whether they overlap or replace each other in these ecological dimensions.

DISCUSSION

Zorapterans have been seen rarely by most entomologists and the few collection records have led to the assumption of high endemism and poor powers of dispersal. To account for these "limited dispersal capabilities," past authors have suggested that the flimsy, narrow, pad-
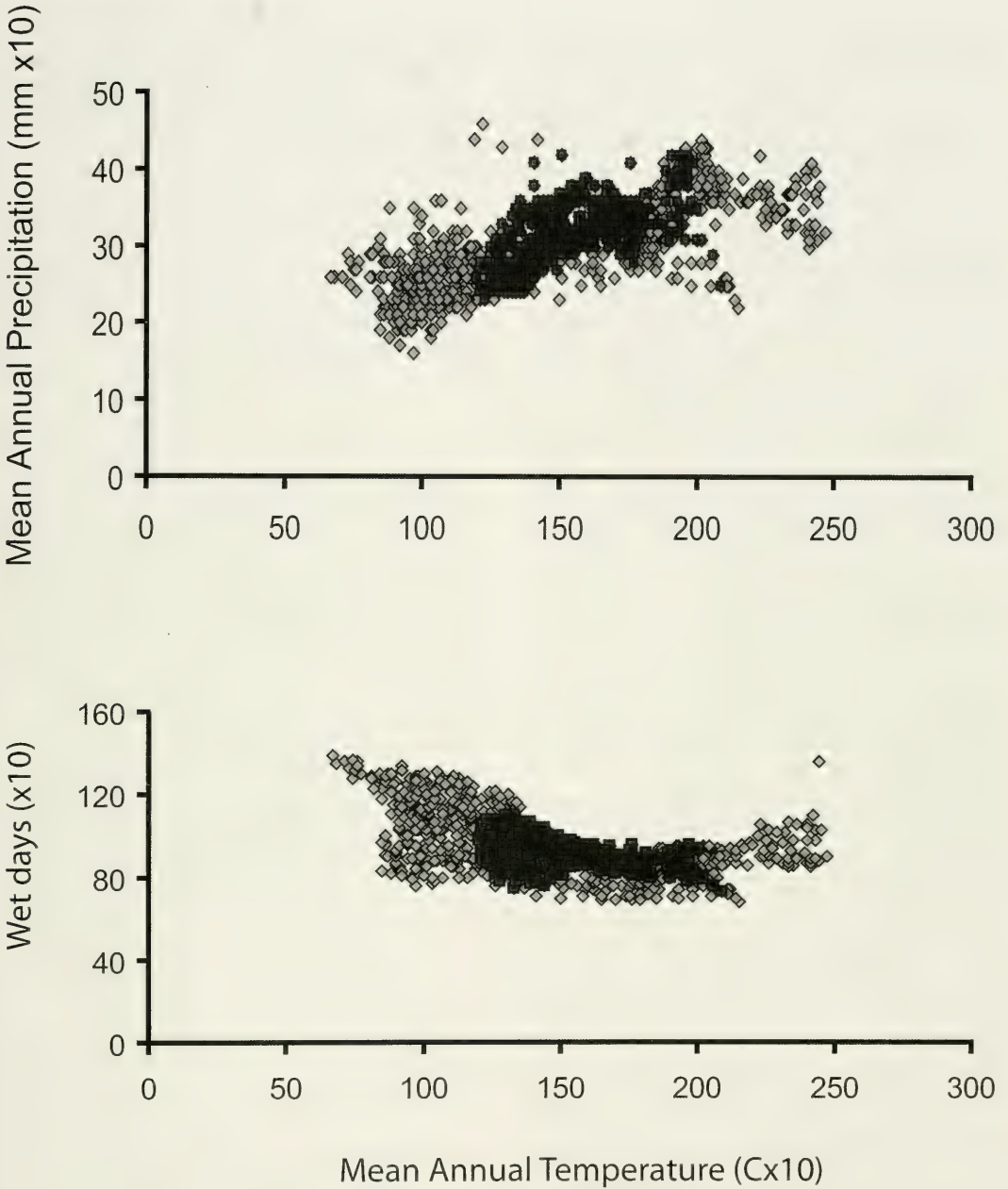
Fig. 3. Visualization of modeled distribution of *Zorotypus hubbardi* in environmental space, viewed in two bivariate plots each of two environmental dimensions. Dark gray diamonds indicate overall availability in the area of analysis; light gray squares indicate the distribution of the species in ecological space relative to availability.

dle-shaped wings of alates preclude effective flight. Indeed, such assumptions have been considered the causal factors for the "narrow" geographical distribu-

tions for most of the species in the order (e.g., Huang 1980). Conversely, the low species diversity within the order and poor sampling of populations has had

more of an impact on the scattered geographical records for zorapterans. This is to say that the "narrow" distributions of most species is likely an artifact of poor sampling (*vide* Engel 2001, 2004). When collectors develop a search image for zorapterans, they tend to find more colonies and broader geographic distributions of species than previously thought (e.g., Engel 2001, 2004). Such evidence indicates that zorapterans may be more mobile than entomologists have appreciated. The wings of zorapterans are no less adept than those of many minute insects (e.g., thrips, microhymenoptera) already known to be capable dispersers. Furthermore, owing to their small sizes, zorapterans can be captured in wind currents and carried great distances. Thus, there are no *a priori* reasons to assume that zorapterans are incapable of regular dispersal.

Zorotypus hubbardi is the best known species of zorapterans in terms of biology and also the one with the greatest number of locality records. Its potential distribution shows that the ecological conditions suitable for the natural occurrence of the species are present in a widespread area, and suggest that zorapteran dispersal capabilities have been underestimated. For Z. *hubbardi*, in particular, the northern and western boundaries of the potential distribution produced by our model, show that occurrence of the species in those areas is feasible given the ecological variables analyzed in this study, and indicate that human activity would not be the only source for the presence of colonies. As suggested by Engel (2003a), zorapterans are likely quite mobile and, at the extremes of zorapteran distribution, colonies are being established regularly in favorable seasons and extirpated during harsh winters. Indeed, populations may be highly mobile and spreading regularly throughout areas of suitable habitability, thereby maintaining specific integrity

over large areas. With global climate change and suitable human intervention (e.g., sawdust piles near homes or lumber yards for artificial warmth during winters), populations gradually push the extremes of their environmental limits. New records of feral colonies (i.e., those apart from human habitation and artificial heat sources) from the areas shown in the potential distribution, particularly the extremes of the range, would provide greater support for our model.

Literature Cited

Anderson, R. P., D. Lew, and A. T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. Ecological Modeling 162: 211–232.

Engel, M. S. 2001. New neotropical records for three *Zorotypus* species (Zoraptera: Zorotypidae). Entomological News 112: 278–280.

———. 2003a. Zoraptera, pp. 1201–1203. *In* Resh, V. H. and R. T. Cardé, eds. Encyclopedia of Insects. Academic Press, San Diego.

———. 2003b. Phylogeny of the Zoraptera. Entomologische Abhandlungen 61: 147–148.

———. 2004. Zoraptera, pp. 637–640. *In* Llorente-Bousquets, J., J. J. Morrone, O. Yáñez-Ordóñez, and I. Fernandez-Vargas, eds. Biodiversidad, Taxonomía y Biogeografía de Artrópodos de México: Hacia una Síntesis de su Conocimiento, Volumen IV. UNAM, México, D.F.

Engel, M. S. and D. A. Grimaldi. 2000. A winged *Zorotypus* in Miocene amber from the Dominican Republic (Zoraptera: Zorotypidae), with discussion on relationships of and within the order. Acta Geologica Hispanica 35: 149–164.

Environmental Systems Research Institute (ESRI), Inc. 2002. ArcView 3.3. 380 New York St., Redlands.

Feria, T. P. and A. T. Peterson. 2002. Using point occurrence data and inferential algorithms to predict local communities of birds. Diversity and Distributions 8: 49–56.

Grimaldi, D. and M. S. Engel. 2005. Evolution of the Insects. Cambridge University Press, Cambridge.

Grinnell, J. 1917. Field tests of theories concerning distributional control. American Naturalist 51: 115–128.

Huang, F.-S. 1980. The order Zoraptera and its geographical distribution, pp. 1087–1090. *In* Liu, D.-S., ed. Geological and Ecological Studies of Qinghai-Xizang Plateau: Proceedings of Symposium on Quinghai-Xizang (Tibet) Plateau. Science Press, Beijing.

MacArthur, R. H. 1972. Geographical Ecology: Patterns in the Distribution of Species. Harper & Row, New York.

Peterson, A. T. 2001. Predicting species' geographic distributions based on ecological niche modeling. Condor 103: 599–605.

Peterson, A. T. and K. C. Cohoon. 1999. Sensitivity of distributional prediction algorithms to geographic data completeness. Ecological Modeling 117: 159–164.

Peterson, A. T. and J. Shaw. 2003. *Lutzomyia* vectors for cutaneous leishmaniasis in southern Brazil: Ecological niche models, predicted geographic distributions, and climate change effects. International Journal for Parasitology 33: 919–931.

Peterson, A. T. and D. A. Vieglais. 2001. Predicting species invasions using ecological niche modeling. BioScience 51: 363–371.

Peterson, A. T., M. Papes, and D. A. Kluza. 2003. Predicting the potential invasive distributions of four alien plant species in North America. Weed Science 51: 863–868.

Rice, N. H., E. Martínez-Meyer, and A. T. Peterson. 2003. Ecological niche differentiation in the *Aphelocoma* jays: A phylogenetic perspective. Biological Journal of the Linnaean Society 80: 369–383.

Rios, N. and H. Bart. 2005. GEOLocate Georeferencing Software for Natural History Collections. URL http://www.museum.tulane.edu/geolocate

Stockwell, D. R. B. 1999. Genetic algorithms II, pp. 123–144. *In* Fielding, A. H., ed. Machine Learning Methods for Ecological Applications. Kluwer Academic Publishers, Boston.

Stockwell, D. R. B. and I. R. Noble. 1992. Induction of sets of rules from animal distribution data: A robust and informative method of analysis. Mathematics and Computers in Simulation 33: 385–390.

Stockwell, D. R. B. and D. P. Peters. 1999. The GARP modeling system: Problems and solutions to automated spatial prediction. International Journal of Geographic Information Systems 13: 143–158.

Stockwell, D. R. B. and A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. Ecological Modeling 148: 1–13.

———. 2003. Comparison of resolution of methods used in mapping biodiversity patterns from point occurrence data. Ecological Indicators 3: 213–221.